2018-01-31

Meta-analysis of health-related quality of life after minimally invasive versus open oesophagectomy for oesophageal cancer

Kauppila, Joonas H; Xie, Shao-Hua; Johar, Asif; Markar, Sheraz; Lagergren, Pernilla

Br J Surg. 2017 Aug;104(9):1131-1140.
http://doi.org/10.1002/bjs.10577
http://hdl.handle.net/10616/46209

Meta-analysis of health-related quality of life after minimally invasive versus open oesophagectomy for oesophageal cancer.

Kauppila, Joonas Henrikki; Xie, Shaohua; Johar, Asif; Markar, Sheraz; Lagergren, Pernilla

DOI: 10.1002/bjs.10577

Title: Health-related quality of life after minimally invasive versus open oesophagectomy for oesophageal cancer - a systematic review and a meta-analysis

Authors: Joonas H Kauppila, MD, PhD [1,2,3], Shaohua Xie, MD, PhD[1], Asif Johar, BSc Hons., MSc[4], Sheraz R Markar, MRCS[1,5], Pernilla Lagergren, RN, PhD.[4]

Affiliations: [1]Upper Gastrointestinal Surgery, Department of Molecular Medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, 17176 Stockholm, Sweden; [2]Department of Surgery and Medical Research Center Oulu, University of Oulu, P.O. Box 5000, 90014 Oulu, Finland; [3]Oulu University Hospital, P.O. Box 21, 90029 Oulu, Finland; [4]Surgical Care Sciences, Department of Molecular medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, 17176 Stockholm, Sweden; [5]Department of Surgery & Cancer, Imperial College London, London, England.

Author correspondence and reprint requests to: Joonas H Kauppila, Upper Gastrointestinal Surgery, Department of Molecular Medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, 17176 Stockholm, Sweden Email: joonas.kauppila@ki.se, Tel. +46 8-517 709 83, Fax: +46 8-517 762 80

Sources of support: This study was supported by grants from the Stockholm Cancer Society (PL), Sigrid Jusélius Foundation (JHK), Orion Research Foundation (JHK) and National Institute for Health Research NIHR-CTF-2015-04-09 (SRM). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Abstract

Background: The aim of this systematic review and meta-analysis was to compare health-related quality of life (HRQOL) outcomes between minimally invasive and open oesophagectomy for cancer at different postoperative time points

Methods: A search of PubMed (MEDLINE), Web of Science, EMBASE, Scopus, CINAHL and Cochrane Library was performed, yielding 2853 titles. Nine studies comparing patients who underwent minimally invasive surgery (n=1161) with those who underwent open surgery (n=903) were included in the systematic review based on eligibility criteria. A random-effects meta-analysis was conducted based on eight studies that measured HRQOL scores using the EORTC QLQ-C30 and QLQ-OES18 questionnaires. Mean score differences (MSD) >10 were considered clinically relevant**. P**ooled effects of MSD with 95% confidence intervals (CI) were estimated to assess statistical significance.

Results: Patients who underwent minimally invasive surgery had on average better scores for global quality of life (MSD 11.6, 95% CI 3.8 to 19.4), physical function (MSD 11.9, 95% CI 3.9 to 19.8), fatigue (MSD -13.2, 95% CI -17.6 to -8.8), and pain (MSD -15.9, 95% CI -20.5 to -11.2) than those who underwent open surgery at 3 months after surgery. At 6 and 12 months after surgery, no significant differences remained.

Conclusion: This meta-analysis indicates **better average global** quality of life, physical function, fatigue and pain 3 months following minimally invasive surgery compared to open surgery. No such differences remain at longer follow-up of 6 and 12 months. This knowledge can guide clinical decision-making and improve patient information.

Introduction

Oesophageal cancer is the 9th most common cancer and the 6th most common cause of cancer death worldwide.[1] The mainstay curative option for most patients with locally advanced cancer is oesophagectomy, often preceded by neoadjuvant therapy.[2] Surgical resection is associated with a high rate of post-operative morbidity, and health-related quality of life (HRQOL) can deteriorate substantially following such radical surgery.[3] In recent years minimally invasive techniques using a combination of thoracoscopic and/or laparoscopic approaches have been used increasingly, with the primary aims of reducing surgical trauma, minimising complications and improving postoperative recovery.[4] In addition to a few cohort studies, one randomized clinical trial has indicated that specific components of short-term HRQOL may be improved through utilization of a minimally invasive approach to esophagectomy.[5] However, great variability exists regarding which HRQOL outcomes have been assessed and the time period after surgery that these were evaluated.[6, 7] There is presently no conclusive evidence as to whether HRQOL in general is improved following minimally invasive oesophagectomy compared to open surgery, if any specific HRQOL outcomes differ between these approaches and the length of time these persist following surgery.

Therefore, the aim of this systematic review and meta-analysis was to clarify potential differences in postoperative HRQOL outcomes over time after minimally invasive compared to open oesophagectomy for oesophageal cancer.

Methods

This systematic review was conducted according to the PRISMA guidelines.[8] An a priori established detailed study protocol was followed.

Search strategy

A comprehensive literature search was conducted in September 2016 using a keyword search on PubMed (MEDLINE), Web of Science, EMBASE, Scopus, CINAHL and Cochrane, from the inception of each database. The following search string was used: (mini-invasiv* OR minimally-invasive OR minimally invasive OR vats OR thoraco* OR laparo*) AND (esophag* OR oesophag* OR gastro-oesophag* OR gastroesophag*) AND (neoplas* OR tumo* OR cancer OR carcinoma OR adenocarcinoma) AND (Quality of life OR qol OR hqol OR hrqol).

Study selection

Studies considered for inclusion in the systematic review and meta-analysis had the following characteristics:

1) Original studies.
2) Written in English.
3) Oesophageal cancer or high-grade dysplasia (HGD) were primary indications for oesophagectomy.
4) HRQOL data were assessed using any well-validated HRQOL instrument or its derivative.
5) HRQOL outcomes were measured both before and after the oesophagectomy.
6) Post-operative HRQOL was compared between minimally invasive and open surgical procedures.

The search identified 2853 titles. The titles and abstracts of these studies were evaluated. When the studies seemed to meet the eligibility criteria, or when information was insufficient to exclude them, the full articles were reviewed. The reference lists of the retrieved articles, PubMed "related articles" and articles dealing with the literature review were scanned for potential additional studies. After the search, one author (J.H.K.) performed the screening of all titles and abstracts. Studies were included according to the eligibility criteria above. After the initial screening, full text articles were obtained. Full-text articles were independently studied by two authors (J.H.K., and P.L). In the case of discrepancies, the studies were discussed based on the pre-determined eligibility criteria between the authors and resulted in a consensus decision. The workflow is summarized in Figure 1.

Data extraction

All data from eligible studies were independently extracted by two investigators (J.H.K., and S.X.) to maintain the integrity of the data. Gathered information included the name of first author, publication year, study period, design, population characteristics (age and sex), treatment, number of patients in the two surgery groups, HRQOL indicators used and HRQOL outcomes at baseline and different time points after surgery. First authors and corresponding authors in the eligible studies were contacted by email up to three times to obtain unreported data.

Quality assessment

Study quality and bias were assessed separately by two authors (J.H.K. and P.L.) using the Newcastle-Ottawa scale, as all but one of the included studies were cohort studies.[9] The pre-determined items for assessment of study quality and results were used to rank the studies. Discrepancies between assessors were settled upon discussion. Bias in individual studies was analysed qualitatively.

Definition of exposure and outcome

The study exposure was divided in two categories, namely minimally invasive oesophagectomy and open oesophagectomy. Minimally invasive oesophagectomy included hybrid approaches such as laparoscopic-assisted, hand video-assisted and thoracoscopic-assisted, and totally minimally invasive (both laparoscopic and thoracoscopic) procedures. Open procedures included Ivor-Lewis- and left transthoracic thoraco-abdominal oesophagectomy (with intra-thoracic anastomosis), three-incision oesophagectomy (with neck anastomosis), and transhiatal oesophagectomy (with neck anastomosis).

Outcome of interest was HRQOL, a multidimensional measure consisting of physical health, psychological health, functional status, social relationships, and personal beliefs.[10] HRQOL can be classified as generic or disease-specific and measured using a variety of scales and indicators. The HRQOL questionnaires that were used in the included studies were the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 v3.0 core questionnaire (EORTC QLQ-C30) measuring cancer-related functions and symptoms,[11] EORTC oesophageal cancer-specific HRQOL module (EORTC QLQ-OES18) assessing disease related symptoms and items,[12] short-form health survey (SF-36) measuring general health,[13] the Multi-Dimensional Fatigue Inventory (MFI-20) measuring different aspects of fatigue,[14] a modified Katz Scale and a modified Lawton and Brody Scale measuring daily activities.[15] The HRQOL questionnaires used in the studies were checked for validity in the native languages of the countries of origin specified in Table 1.

Statistical analysis

The statistical analyses were performed using RevMan version 5.3 (The Cochrane Collaboration, Oxford, UK) and STATA version 13.0 (StataCorp LP), according to the Cochrane guidelines for systematic reviews.[9] Number of the patients in each treatment group was collected as reported in the individual studies. Standard deviations were calculated based on the standard error of the mean value and patient number at each time point for each of the two surgery groups. Missing data on standard deviations were imputed using the largest reported standard deviation for the outcome with the given surgical operation at a given time point.[9, 16] This approach was used to obtain the most conservative effect estimate. Data available only in graphical format were extracted using the WebPlotDigitizer tool (http://arohatgi.info/WebPlotDigitizer). In one included study,[17] two open oesophagectomy procedures were combined as suggested in the Cochrane Handbook.[9] To evaluate potential bias due to imputation of standard deviations for two of the included studies, we performed a sensitivity analysis including only the studies that did not require any data to be imputed.

Continuous variables were analysed using inverse variance according to the DerSimonian-Laird method.[18] To obtain estimates of the average treatment effects, a random effects model was used to estimate mean differences for continuous data across the studies.[19] Mean score differences (MSD) and pooled effects with 95% confidence intervals (CI) were depicted on forest plots. Based on previous literature, the mean differences were considered clinically significant only if the pooled differences were at least 10 points, which correspond to at least a "moderate" change or difference for the patients.[20, 21] Statistical heterogeneity of the studies was assessed in terms of the I^2 statistic.[22] An I^2 statistic <25% indicated a minor inconsistency, and an I^2 statistic >50% indicated a major inconsistency. As the random-effects meta-analysis calculates the average effect of a given treatment, 95% prediction intervals were calculated for each outcome at a given time point to estimate the range of true difference between the treatments in 95% of the population.[19, 23] As the number of included studies was small, meta-regression could not be reliably performed.[24] Thus, a sensitivity analysis was undertaken by removing one study at a time in the meta-analysis to estimate the influence of each study on the pooled estimate of HRQOL scores. Potential publication bias and small-study effects of the clinically relevant HRQOL outcomes were analysed by visually inspecting the funnel plots instead of statistical testing, given the small number of included studies.[25]

Results

Included studies

Among all 2853 titles, 9 articles published from 2010 to 2016 were eligible for qualitative analysis (Figure 1).[5-7, 15, 17, 26-29] Main features of these studies are summarized in Table 1. In total, the 9 studies included 2164 patients, of whom 1161 (58.3%) underwent minimally invasive surgery and 903 (41.7%) underwent open surgery. The median number of patients per study was 114 (range 56-888). In three studies, only patients with oesophageal squamous cell carcinoma were included,[17, 28, 29] while one study included only patients with adenocarcinoma.[26] Five studies also included patients with cancer of the gastroesophageal junction,[5-7, 15, 27] and **two** stud**ies** included patients with high-grade dysplasia.[7, 15] The presence of medical comorbidities was not described in two of the studies,[15, 27] whereas fitness evaluations with American Society of Anesthesiologists grades (or ASA grades) were available in four studies.[5, 6, 26, 29] In the 7 studies that described the use of neoadjuvant therapy, 595 (32.4%) of the 1836 patients received such treatment.[5, 6, 15, 17, 26, 27, 29] The number of patients receiving postoperative chemo- or chemoradiotherapy was described in only one of the 9 studies.[26]

The main open surgical procedure was the right-sided transthoracic (Ivor-Lewis) oesophagectomy. Less frequent procedures included McKeown modification in an undefined number of patients, as well as left thoracoabdominal and left transthoracic oesophagectomy. No patients undergoing transhiatal resections were included in the studies. Thus, all patients in the open surgery group underwent thoracotomy. There was variability in the anastomotic technique and location. Of patients with minimally invasive surgery, at least 249 underwent totally minimally invasive oesophagectomy in five studies.[5, 7, 15, 17, 26] Five studies had a hybrid approach or included both totally minimally invasive and hybrid minimally invasive oesophagectomy in the treatment group (912 patients).[6, 15, 27-29] One of the studies included a comparison between laparoscopy-thoracoscopy approach and hybrid minimally invasive procedures including laparotomy or thoracotomy.[15] One study utilized hand video-assisted thoracoscopic surgery as the minimally invasive procedure.[28] Taken together, all studies included a minimally invasive group of patients who underwent thoracoscopy and an open oesophagectomy group of who underwent thoracotomy. Details on patients and surgery are further described in Tables 2 and 3.

The HRQOL outcomes were measured using both the EORTC QLQ-C30 and the QLQ-OES18 questionnaires in eight studies.[5-7, 17, 26-29] In one study, MFI-20 and activities of daily living questionnaires suggested decreased fatigue and increased proportion of independent patients after minimally invasive surgery compared to open surgery at 3 and 6 months.[15] All studies using the EORTC questionnaires measured HRQOL at baseline and postoperatively. One of the studies reported one-week outcomes, three studies reported two-week outcomes, four studies reported four-week outcomes, and two studies reported six-week outcomes. Seven of the studies measured HRQOL outcomes at 3-months and six at 6-months after surgery. Four studies measured 12-month HRQOL outcomes and two studies also reported the 24-month outcomes. Reporting on HRQOL outcomes was generally selective; only three articles reported all or almost all of the numerical values for the EORTC QLQ-C30 and QLQ-OES18 questionnaire outcomes. Two studies did not present any numerical values, but instead provided graphs. Only one study used evidence-based cut-off values for interpreting the differences in the HRQOL scores.[6, 20, 21]

General cancer-related symptoms (EORTC QLQ-C30)

One study measuring fatigue with MFI-20 scale was excluded from the quantitative analysis.[15] The study did not report general fatigue and the subscales were not compatible with fatigue scale in the EORTC questionnaire. The meta-analysis was conducted on 4-6 week and 3-, 6- and 12-month extracted outcome data on global quality of life, physical function, fatigue, and pain from the 8 included studies. Due to the small number of studies reporting outcomes at time points less than 6 weeks postoperatively, outcomes at 4-6 weeks from 5 studies were combined and analysed. Clinical heterogeneity was found between the study populations, as there were differences between the sub-location and histology of the cancers, as well as the study design. The I^2 statistic indicated high statistical heterogeneity (>50%) for all of the analyses, except the global quality of life at 4-6 weeks (I^2=25%) and 12 months (I^2=9%), fatigue at 12 months (I^2=0%) and pain at 4-6 weeks (I^2=25%). However, the baseline values for the studied outcomes in each study were similar in the treatment groups, with no MSD greater than 5 points between surgery groups.

The meta-analyses of these key HRQOL outcomes showed that minimally invasive surgery was followed by better outcomes at 4-6 weeks and 3 months compared to open surgery. The MSDs at 4-6 weeks were 16.1, 95% CI 13.9 to 18.3 for global quality of life, 26.9, 95% CI 19.0 to 34.9 for physical function, -18.8, 95% CI -29.3 to -8.3 for fatigue and -29.0, 95% CI -31.6 to -26.5 for pain. The 3 month outcomes were also better after minimally invasive oesophagectomy for global quality of life (MSD 11.6, 95% CI 3.8 to 19.4), physical function (MSD 11.9, 95% CI 3.9 to 19.8), fatigue (MSD -13.2, 95% CI -17.6 to -8.8), and pain (MSD -15.9, 95% CI -20.5 to -11.2). After 3 months there were no clinically significant MSDs in HRQOL outcomes, except for physical function at 6 months in favor of minimally invasive surgery (MSD 11.8, 95% CI -0.4 to 24.0), but this difference was not statistically significant. The forest plots and pooled estimates for the four outcomes are shown at 4-6 weeks

(Supplementary figure 1), 3 months (Figure 2), 6 months (Supplementary figure 2) and 12 months (Supplementary figure 3). Figure 3 shows the pooled differences of these key outcomes at different time points with 95 CIs. Majority of the patients should experience better global HRQOL and less pain at 4-6 weeks and less fatigue and pain at 3 months after surgery, as reflected by the 95% prediction intervals (Suppl. Figure 4).

The sensitivity analysis of the pooled results excluding the studies using some imputed data showed that the observed pooled clinically significant differences at 3 months remained. The sensitivity analysis omitting one study at a time showed that the omission of one of the studies[29] made the 3-month difference of surgical method on global quality of life and physical function clinically less relevant, but the difference remained statistically significant (MSD 8.5, 95% CI 5.6 to 11.4 for global HRQOL and MSD 9.0, 95% CI 4.7 to 13.4 for physical function, at 3 months). Sensitivity analysis was also conducted excluding the studies with hybrid minimally invasive operations. Based on pooled results from two to four studies at each time point, totally minimally invasive surgery had clinically (MSD >10 points) and statistically significant superior outcomes compared to open surgery until 3 months, but not at 6 months, similarly to the main analysis. At 12 months, the difference in global HRQOL was not clinically significant. Other key outcomes at 12 months could not be analysed as data was available from only one study.

Two studies showed decreased dyspnoea symptoms at 4-6 weeks after surgery in favour of minimally invasive esophagectomy.[27, 29] There were no differences in other function scales (role, emotional, cognitive, or social) or symptom scales and items (nausea and vomiting, insomnia, appetite loss, constipation, diarrhoea, or financial problems) between the groups.

However, the lack of reporting and obtaining the numerical values for these outcomes prevented formal meta-analysis of these outcomes.

Oesophageal cancer-related symptoms (QLQ-OES18)

There were no clinically relevant differences between open and minimally invasive oesophagectomy for most of the QLQ-OES18 outcomes at any time points in any of the studies. Six of the eight studies did not report the numerical data for the outcomes for one or more items. Thus, unbiased meta-analysis of these outcomes could not be performed. One of the studies[7] reported statistically significant, but not clinically relevant, less oesophageal pain at 3 months, and two studies reported clinically relevant less oesophageal pain at 6 weeks[5] and 12 months[5, 6] after minimally invasive surgery compared to open surgery. Also, two studies[5, 7] reported clinically relevant fewer speech problems at 6-12 weeks postoperatively after minimally invasive surgery compared to open oesophagectomy.

Discussion

The results indicate that minimally invasive surgery is on average followed by better postoperative outcomes regarding global quality of life, physical function, fatigue, and pain up to 3 months after surgery, but these differences fail to persist at 6 or 12 months postoperatively.

Methodological advantages of this meta-analysis include the strict inclusion criteria, use of well-validated questionnaires in all of the included studies, and using a cut-off of 10 for MSDs to reduce bias from multiple testing. [11, 12, 20, 21] The affected components between the surgical approaches had the same direction of effect at 4-6 weeks and 3 months postoperatively in all studies, suggesting actual effect of the treatment. Additionally, sensitivity analyses by removing any single study, all studies with extracted data or all studies that had hybrid minimally invasive operations from the analysis did not have major effect on the conclusions. Biases in the individual studies might affect the observed pooled effect size. Visual inspection of the funnel plots revealed no evidence of major small-study effects or publication bias. The effect estimates for clinically significant outcomes in the largest study in the meta-analysis were larger than the average effect estimate.[29] Thus, no adjustment using the trim and fill method could be reliably done, as suggested earlier. [30, 31]

The surgical techniques used were variable in both open and minimally invasive groups. All included studies had however a thoracotomy group and a thoracoscopy group. Patients undergoing open surgery had variable procedures, such as Ivor-lewis and left-sided transthoracic oesophagectomy.[17] These operations might have different postoperative recovery profiles. Studies included in the 4-6 weeks outcome analysis were only from Eastern countries, and might not be generalizable for this time point. The results were similar at later time points with Western studies included. Selective reporting and the inability to obtain the missing data may cause information bias. A weakness was the inability to perform meta-regression and adjust the HRQOL outcomes for confounding, such as neoadjuvant treatment, due to the small number of studies.[32] Patients selected for minimally invasive surgery might also be healthier than those selected for open surgery.[7] Preoperative comorbidities increase complications, as well as poor quality of life outcomes.[33, 34] Similar in-study preoperative baseline HRQOL values between treatment groups should adjust for some of these differences.[35, 36] The 95% prediction intervals were calculated to depict the heterogeneity of the studies. It seems that most patients have better global HRQOL and less fatigue at 3 months or less pain up to 3 months after minimally invasive surgery compared to open surgery. Minimally invasive oesophagectomy might reduce complications, which could be the causal link in the improvements in HRQOL up to 3 months following surgery.

This first meta-analysis comparing HRQOL outcomes between minimally invasive and open oesophagectomy identified benefits in four key outcomes following minimally invasive surgery in the short term. No differences remained with longer-term follow up. Most patients are likely to have some benefit from minimally invasive surgery in less pain up to 3 months and less fatigue and better global HRQOL at 3 months after surgery. There are some clinical and research implications that can be drawn from this meta-analysis. Although the 3-month perspective is relevant for patients and healthcare, the lack of differences in HRQOL after this period indicates that open surgery does not need to be abandoned, particularly if the surgeons are more comfortable with open surgery or in patients with contraindications to a minimally invasive approach. The learning curve can decrease the prognosis both in the short- and long term.[37] This might be a particular concern for minimally invasive surgery, as shown in esophagectomy,[38-40] as well as other procedures.[41-43] Therefore, these results are not enough to generally recommend minimally invasive oesophagectomy. Patients want information about HRQOL outcomes after cancer surgery, but surgeons rarely inform the patients about such outcomes.[44] Trauma related to thoracotomy is associated with postoperative pain.[45, 46] Pain and decreased exertion caused by open surgery may severely influence respiratory function, pulmonary complication rate and postoperative HRQOL.[47-49] The present study can help surgeons to inform patients about average effects on HRQOL when making treatment decisions. On-going and future studies and meta-analyses will increase information on the effect of minimally invasive surgery on postoperative HRQOL. Reporting all measured HRQOL outcomes and evaluating the relationship of complications and HRQOL outcomes is important in future studies.

In conclusion, minimally invasive oesophagectomy with thoracoscopy seems to have superior HRQOL outcomes up to 3 months postoperatively compared to open surgery, but no differences remained after this initial postoperative period. These findings cannot be used for any changes in general recommendations on surgical approach, but can help inform the patient about average expected outcomes following oesophageal surgery.

Acknowledgements

This study was supported by grants from the Stockholm Cancer Society (PL), Sigrid Jusélius Foundation (JHK), Orion Research Foundation (JHK) and National Institute for Health Research NIHR-CTF-2015-04-09 (SRM). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. The authors state no conflicts of interest.

References

1. Global Burden of Disease Cancer C, Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, Allen C, Hansen G, Woodbrook R, Wolfe C, Hamadeh RR, Moore A, Werdecker A, Gessner BD, Te Ao B, McMahon B, Karimkhani C, Yu C, Cooke GS, Schwebel DC, Carpenter DO, Pereira DM, Nash D, Kazi DS, De Leo D, Plass D, Ukwaja KN, Thurston GD, Yun Jin K, Simard EP, Mills E, Park EK, Catala-Lopez F, deVeber G, Gotay C, Khan G, Hosgood HD, 3rd, Santos IS, Leasher JL, Singh J, Leigh J, Jonas JB, Sanabria J, Beardsley J, Jacobsen KH, Takahashi K, Franklin RC, Ronfani L, Montico M, Naldi L, Tonelli M, Geleijnse J, Petzold M, Shrime MG, Younis M, Yonemoto N, Breitborde N, Yip P, Pourmalek F, Lotufo PA, Esteghamati A, Hankey GJ, Ali R, Lunevicius R, Malekzadeh R, Dellavalle R, Weintraub R, Lucas R, Hay R, Rojas-Rueda D, Westerman R, Sepanlou SG, Nolte S, Patten S, Weichenthal S, Abera SF, Fereshtehnejad SM, Shiue I, Driscoll T, Vasankari T, Alsharif U, Rahimi-Movaghar V, Vlassov VV, Marcenes WS, Mekonnen W, Melaku YA, Yano Y, Artaman A, Campos I, MacLachlan J, Mueller U, Kim D, Trillini M, Eshrati B, Williams HC, Shibuya K, Dandona R, Murthy K, Cowie B, Amare AT, Antonio CA, Castaneda-Orjuela C, van Gool CH, Violante F, Oh IH, Deribe K, Soreide K, Knibbs L, Kereselidze M, Green M, Cardenas R, Roy N, Tillmann T, Li Y, Krueger H, Monasta L, Dey S, Sheikhbahaei S, Hafezi-Nejad N, Kumar GA, Sreeramareddy CT, Dandona L, Wang H, Vollset SE, Mokdad A, Salomon JA, Lozano R, Vos T, Forouzanfar M, Lopez A, Murray C, Naghavi M. The Global Burden of Cancer 2013. *JAMA Oncol* 2015;**1**(4): 505-527.
2. Kauppila JH, Lagergren J. The surgical management of esophago-gastric junctional cancer. *Surg Oncol* 2016;**25**(4): 394-400.
3. Jacobs M, Macefield RC, Elbers RG, Sitnikova K, Korfage IJ, Smets EM, Henselmans I, van Berge Henegouwen MI, de Haes JC, Blazeby JM, Sprangers MA. Meta-analysis shows clinically relevant and long-lasting deterioration in health-related quality of life after esophageal cancer surgery. *Qual Life Res* 2014;**23**(4): 1155-1176.
4. Biere SS, van Berge Henegouwen MI, Maas KW, Bonavina L, Rosman C, Garcia JR, Gisbertz SS, Klinkenbijl JH, Hollmann MW, de Lange ES, Bonjer HJ, van der Peet DL, Cuesta MA. Minimally invasive versus open oesophagectomy for patients with oesophageal cancer: a multicentre, open-label, randomised controlled trial. *Lancet* 2012;**379**(9829): 1887-1892.
5. Maas KW, Cuesta MA, van Berge Henegouwen MI, Roig J, Bonavina L, Rosman C, Gisbertz SS, Biere SS, van der Peet DL, Klinkenbijl JH, Hollmann MW, de Lange ES, Bonjer HJ. Quality of Life and Late Complications After Minimally Invasive Compared to Open Esophagectomy: Results of a Randomized Trial. *World J Surg* 2015;**39**(8): 1986-1993.
6. Barbour AP, Mc Cormack OM, Baker PJ, Hirst J, Krause L, Brosda S, Thomas JM, Blazeby JM, Thomson IG, Gotley DC, Smithers BM. Long-Term Health-related Quality of Life Following Esophagectomy: A Nonrandomized Comparison of Thoracoscopically Assisted and Open Surgery. *Ann Surg* 2016.
7. Nafteux P, Moons J, Coosemans W, Decaluwe H, Decker G, De Leyn P, Van Raemdonck D, Lerut T. Minimally invasive oesophagectomy: a valuable alternative to open oesophagectomy for the treatment of early oesophageal and gastro-oesophageal

junction carcinoma. *Eur J Cardiothorac Surg* 2011;**40**(6): 1455-1463; discussion 1463-1454.
8. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**: b2535.
9. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. *The Cochrane Collaboration* 2011.
10. Hennessy CH, Moriarty DG, Zack MM, Scherr PA, Brackbill R. Measuring health-related quality of life for public health surveillance. *Public Health Rep* 1994;**109**(5): 665-672.
11. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**(5): 365-376.
12. Blazeby JM, Conroy T, Hammerlid E, Fayers P, Sezer O, Koller M, Arraras J, Bottomley A, Vickery CW, Etienne PL, Alderson D, European Organisation for R, Treatement of Cancer G, Quality of Life G. Clinical and psychometric validation of an EORTC questionnaire module, the EORTC QLQ-OES18, to assess quality of life in patients with oesophageal cancer. *Eur J Cancer* 2003;**39**(10): 1384-1394.
13. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**(6): 473-483.
14. Smets EM, Garssen B, Cull A, de Haes JC. Application of the multidimensional fatigue inventory (MFI-20) in cancer patients receiving radiotherapy. *Br J Cancer* 1996;**73**(2): 241-245.
15. Parameswaran R, Titcomb DR, Blencowe NS, Berrisford RG, Wajed SA, Streets CG, Hollowood AD, Krysztopik R, Barham CP, Blazeby JM. Assessment and comparison of recovery after open and minimally invasive esophagectomy for cancer: an exploratory study in two centers. *Ann Surg Oncol* 2013;**20**(6): 1970-1977.
16. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006;**59**(1): 7-10.
17. Zeng J, Liu JS. Quality of life after three kinds of esophagectomy for cancer. *World J Gastroenterol* 2012;**18**(36): 5106-5113.
18. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**(3): 177-188.
19. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;**342**: d549.
20. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res* 1996;**5**(6): 555-567.
21. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998;**16**(1): 139-144.
22. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414): 557-560.
23. Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Res Synth Methods* 2017.
24. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002;**21**(11): 1559-1573.

25. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;**53**(11): 1119-1129.
26. Hong L, Zhang Y, Zhang H, Yang J, Zhao Q. The short-term outcome of three-field minimally invasive esophagectomy for Siewert type I esophagogastric junctional adenocarcinoma. *Ann Thorac Surg* 2013;**96**(5): 1826-1831.
27. Wang H, Feng M, Tan L, Wang Q. Comparison of the short-term quality of life in patients with esophageal cancer after subtotal esophagectomy via video-assisted thoracoscopic or open surgery. *Dis Esophagus* 2010;**23**(5): 408-414.
28. Shen H, Wang J, Li W, Yi W, Wang W. Assessment of health-related quality of life of patients with esophageal squamous cell carcinoma following esophagectomy using EORTC quality of life questionnaires. *Mol Clin Oncol* 2015;**3**(1): 133-138.
29. Wang H, Shen Y, Feng M, Zhang Y, Jiang W, Xu S, Tan L, Wang Q. Outcomes, quality of life, and survival after esophagectomy for squamous cell carcinoma: A propensity score-matched comparison of operative approaches. *J Thorac Cardiovasc Surg* 2015;**149**(4): 1006-1014; discussion 1014- 1005 e1004.
30. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat Med* 2003;**22**(13): 2113-2126.
31. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Stat Med* 2007;**26**(25): 4544-4562.
32. Safieddine N, Xu W, Quadri SM, Knox JJ, Hornby J, Sulman J, Wong R, Guindi M, Keshavjee S, Darling G. Health-related quality of life in esophageal cancer: effect of neoadjuvant chemoradiotherapy followed by surgical intervention. *J Thorac Cardiovasc Surg* 2009;**137**(1): 36-42.
33. Djarv T, Blazeby JM, Lagergren P. Predictors of postoperative quality of life after esophagectomy for cancer. *J Clin Oncol* 2009;**27**(12): 1963-1968.
34. Backemar L, Lagergren P, Djarv T, Johar A, Wikman A, Lagergren J. Comorbidities and Risk of Complications After Surgery for Esophageal Cancer: A Nationwide Cohort Study in Sweden. *World J Surg* 2015;**39**(9): 2282-2288.
35. Briez N, Piessen G, Bonnetain F, Brigand C, Carrere N, Collet D, Doddoli C, Flamein R, Mabrut JY, Meunier B, Msika S, Perniceni T, Peschaud F, Prudhomme M, Triboulet JP, Mariette C. Open versus laparoscopically-assisted oesophagectomy for cancer: a multicentre randomised controlled phase III trial - the MIRO trial. *BMC Cancer* 2011;**11**: 310.
36. Kataoka K, Takeuchi H, Mizusawa J, Ando M, Tsubosa Y, Koyanagi K, Daiko H, Matsuda S, Nakamura K, Kato K, Kitagawa Y, Japan Esophageal Oncology Group/Japan Clinical Oncology G. A randomized Phase III trial of thoracoscopic versus open esophagectomy for thoracic esophageal cancer: Japan Clinical Oncology Group Study JCOG1409. *Jpn J Clin Oncol* 2016;**46**(2): 174-177.
37. Markar SR, Mackenzie H, Lagergren P, Hanna GB, Lagergren J. Surgical Proficiency Gain and Survival After Esophagectomy for Cancer. *J Clin Oncol* 2016;**34**(13): 1528-1536.
38. Tapias LF, Morse CR. Minimally invasive Ivor Lewis esophagectomy: description of a learning curve. *J Am Coll Surg* 2014;**218**(6): 1130-1140.
39. Wang Q, Wu Z, Chen G, Zhang S, Shen G, Wu M. Two-Stage Indicators to Assess Learning Curves for Minimally Invasive Ivor Lewis Esophagectomy. *Thorac Cardiovasc Surg* 2016.

40. Mackenzie H, Markar SR, Askari A, Ni M, Faiz O, Hanna GB. National proficiency-gain curves for minimally invasive gastrointestinal cancer surgery. *Br J Surg* 2016;**103**(1): 88-96.
41. Park IJ, Choi GS, Lim KH, Kang BM, Jun SH. Multidimensional analysis of the learning curve for laparoscopic colorectal surgery: lessons from 1,000 cases of laparoscopic colorectal surgery. *Surg Endosc* 2009;**23**(4): 839-846.
42. Voitk AJ, Tsao SG, Ignatius S. The tail of the learning curve for laparoscopic cholecystectomy. *Am J Surg* 2001;**182**(3): 250-253.
43. Wang M, Meng L, Cai Y, Li Y, Wang X, Zhang Z, Peng B. Learning Curve for Laparoscopic Pancreaticoduodenectomy: a CUSUM Analysis. *J Gastrointest Surg* 2016;**20**(5): 924-935.
44. McNair AG, MacKichan F, Donovan JL, Brookes ST, Avery KN, Griffin SM, Crosby T, Blazeby JM. What surgeons tell patients and what patients want to know before major cancer surgery: a qualitative study. *BMC Cancer* 2016;**16**: 258.
45. Gottschalk A, Cohen SP, Yang S, Ochroch EA. Preventing and treating pain after thoracic surgery. *Anesthesiology* 2006;**104**(3): 594-600.
46. Elshiekh MA, Lo TT, Shipolini AR, McCormack DJ. Does muscle-sparing thoracotomy as opposed to posterolateral thoracotomy result in better recovery? *Interact Cardiovasc Thorac Surg* 2013;**16**(1): 60-67.
47. Shea RA, Brooks JA, Dayhoff NE, Keck J. Pain intensity and postoperative pulmonary complications among the elderly after abdominal surgery. *Heart Lung* 2002;**31**(6): 440-449.
48. Derogar M, Orsini N, Sadr-Azodi O, Lagergren P. Influence of major postoperative complications on health-related quality of life among long-term survivors of esophageal cancer surgery. *J Clin Oncol* 2012;**30**(14): 1615-1619.
49. Viklund P, Lindblad M, Lagergren J. Influence of surgery-related factors on quality of life after esophageal or cardia cancer resection. *World J Surg* 2005;**29**(7): 841-848.

Figure legends.

Figure 1. Selection of the studies included in the systematic review and meta-analysis.

Figure 2. Results of a meta-analysis of the health-related quality of life (HRQOL) outcomes at 3 months after oesophagectomy for cancer. The mean differences in HRQOL scores are depicted in forest plots, with horizontal lines showing 95% confidence intervals (CI), for global HRQOL (A), physical function (B), fatigue (C) and pain (D). Higher scores in A and B indicate better function and higher scores in C and D indicate worse symptoms. Difference of more than 10 points is considered clinically relevant. IV, inverse variance; MIE, minimally invasive oesophagectomy; OE, open oesophagectomy.

Figure 3. Pooled key health-related quality of life mean score differences (MSD)and 95% confidence intervals (CI) comparing minimally invasive versus open oesophagectomy at different time points. Negative MSDs were converted positive for clarity. Larger MSDs indicate better outcome after minimally invasive compared to open surgery. Dashed horizontal line indicates the threshold for clinical significance of each pooled mean difference.

Supplementary figure 1. Results of a meta-analysis of the health-related quality of life (HRQOL) outcomes at 4-6 weeks after oesophagectomy for cancer. The mean differences in HRQOL scores are depicted in forest plots, with horizontal lines showing 95% confidence intervals (CI), for global HRQOL (A), physical function (B), fatigue (C) and pain (D). Higher scores in A and B indicate better function and higher scores in C and D indicate worse symptoms. Difference of more than 10 points is considered clinically relevant. Number of patients in Maas et al was obtained from Biere et al.[4] IV, inverse variance; MIE, minimally invasive oesophagectomy; OE, open oesophagectomy

Supplementary figure 2. Results of a meta-analysis of the health-related quality of life (HRQOL) outcomes at 6 months after oesophagectomy for cancer. The mean differences in HRQOL scores are depicted in forest plots, with horizontal lines showing 95% confidence intervals (CI), for global HRQOL (A), physical function (B), fatigue (C) and pain (D). Higher scores in A and B indicate better function and higher scores in C and D indicate worse symptoms. Difference of more than 10 points is considered clinically relevant. IV, inverse variance; MIE, minimally invasive oesophagectomy; OE, open oesophagectomy.

Supplementary figure 3. Results of a meta-analysis of the health-related quality of life (HRQOL) outcomes at 12 months after oesophagectomy for cancer. The mean differences in HRQOL scores are depicted in forest plots, with horizontal lines showing 95% confidence intervals (CI), for global HRQOL (A), physical function (B), fatigue (C) and pain (D). Higher scores in A and B indicate better function and higher scores in C and D indicate worse symptoms. Difference of more than 10 points is considered clinically relevant. IV, inverse variance; MIE, minimally invasive oesophagectomy; OE, open oesophagectomy.

Supplementary figure 4. Pooled key health-related quality of life mean score differences (MSD) and 95% prediction intervals comparing minimally invasive versus open oesophagectomy for cancer. Negative MSDs were converted positive for clarity. Larger MSDs indicate better outcome after minimally invasive compared to open surgery. Dashed horizontal line indicates the threshold for clinical significance of each pooled mean difference.

Tables

Table 1. Characteristics, quality and HRQOL assessment of the 9 studies included in the systematic review.

Reference	Study interval	Country	Study mode	Study quality	Number of patients	HRQOL questionnaires	Questionnaire compliance	Included time points in months
Barbour, 2016	1998-2011	Australia	Cohort	8	487	QLQ-C30, QLQ-OES18	>95%	3, 6, 12
Hong, 2013	2009-2012	China	Randomized Cohort	8	114	QLQ-C30, QLQ-OES18	93%	1, 3
Maas, 2015	2009-2011	Italy, Netherlands, Spain	RCT	8	115	QLQ-C30, QLQ-OES18, SF-36	82%	1.5, 12
Nafteux, 2011	2005-2010	Belgium	Cohort	7	166	QLQ-C30, QLQ-OES18	86,5%	3, 6, 12
Parameswaran, 2013	2007-2008	U.K.	Cohort	7	86	MFI-20, Brody-lawton, Katz scale	77%	1.5, 3, 6
Shen, 2015	2005-2007	China	Randomized cohort	7	62	QLQ-C30, QLQ-OES18	95%	1, 3, 6
Wang, 2010	2007-2008	China	Cohort	7	56	QLQ-C30, QLQ-OES18	95%	1, 3, 6
Wang, 2015	2004-2013	China	Cohort	7	888	QLQ-C30, QLQ-OES18	100%	1, 3, 6, 12
Zeng, 2012	2010	China	Cohort	8	90	QLQ-C30, QLQ-OES18	90%	3, 6

Abbreviations: QLQ-C30, European Organisation for Cancer Research and Treatment general quality of life questionnaire Core 30; QLQ-OES18, European Organisation for Cancer Research and Treatment oesophageal cancer specific quality of life questionnaire Oesophageal 18; RCT, Randomized clinical trial; SF-36, Short Form (36) Health Survey 36, MFI-20, Multi-dimensional Fatigue Inventory

Table 2. Surgery, gender age and neoadjuvant therapy distributions of the patients in the included studies

References	Number of patients			Gender Male/Female		Mean age		Neoadjuvant therapy	
	Open	**HMIO**	**TMIO**	**Open**	**MIO**	**Open**	**MIO**	**Open**	**MIO**
Barbour, 2016	110	377	0	98/12	316/61	64*	64*	56/110	202/377
Hong, 2013	55	0	59	41/18	38/17	56	56	0/55	0/59
Maas, 2015	56	0	59	46/10	43/16	62*	62*	56/56	59/59
Nafteux, 2011	101	0	65	82/19	49/16	64	63	ND	ND
Parameswaran, 2013	19	31	36	15/4	47/20	64*	67, 64*[a]	17/19	27/31, 23/36
Shen, 2015	29	33	0	23/6	25/8	ND	ND	ND	ND
Wang, 2010	29	27	0	19/10	19/8	58	61	0/29	0/27
Wang, 2015	444	63	381†	358/86	362/82	56*	56*	75/444	80/444
Zeng, 2012	30, 30[b]	0	30	19/11, 20/10[b]	21/9	58, 63[b]	66	0/60	0/30

†analysed as HMIO; *median; [a]Two groups, HMIO, TMIO; [b]Two groups, Ivor-Lewis, left transthoracic
Abbreviations: MIO, minimally invasive oesophagectomy; HMIO, hybrid minimally invasive oesophagectomy; TMIO, totally minimally invasive oesophagectomy; ND, not described.

Table 3. Tumour stage, operative time, rate of R0 resection margins and hospital stay in the included studies.

References	Tumour stage (HGD/0-I/II/III-IV)		Mean operation time (min)		R0 resection margins (%)		Hospital stay in days (median)	
	Open	MIO	Open	MIO	Open	MIO	Open	MIO
Barbour, 2016	0/16/35/59	0/131/125/121	ND	ND	82	91	15	13
Hong, 2013	0/0/24/35	0/0/19/36	362	521	100	100	26	16
Maas, 2015	0/11/22/19	0/14/26/15	299†	329†	84	92	14†	11†
Nafteux, 2011	14/87/0/0	10/55/0/0	322	375	100	100	11	10
Parameswaran, 2013	0/0/8/11	0/6/12/13, 5/7/13/10[a]	330	360, 390[a]	ND	ND	ND	ND
Shen, 2015	0/20/7/2	0/22/10/1	ND	ND	ND	ND	ND	ND
Wang, 2010	0/5/19/5	0/6/18/3	309	267	ND	ND	4*	2*
Wang, 2015	0/69/255/120	0/62/254/128	211	191	ND	ND	12	11
Zeng, 2012	0/2/9/19, 0/3/11/16[b]	0/0/13/17	287, 143[b]	306	ND	ND	ND	ND

[a] two groups, hybrid minimally invasive surgery, totally minimally invasive surgery; [b]two groups; Ivor-Lewis oesophagectomy, left transthoracic oesophagectomy; †data from Biere, 2012[4]; *ICU stay.
Abbreviations: HGD, High-grade dyplasia; MIO, minimally invasive oesophagectomy.

Figure 1

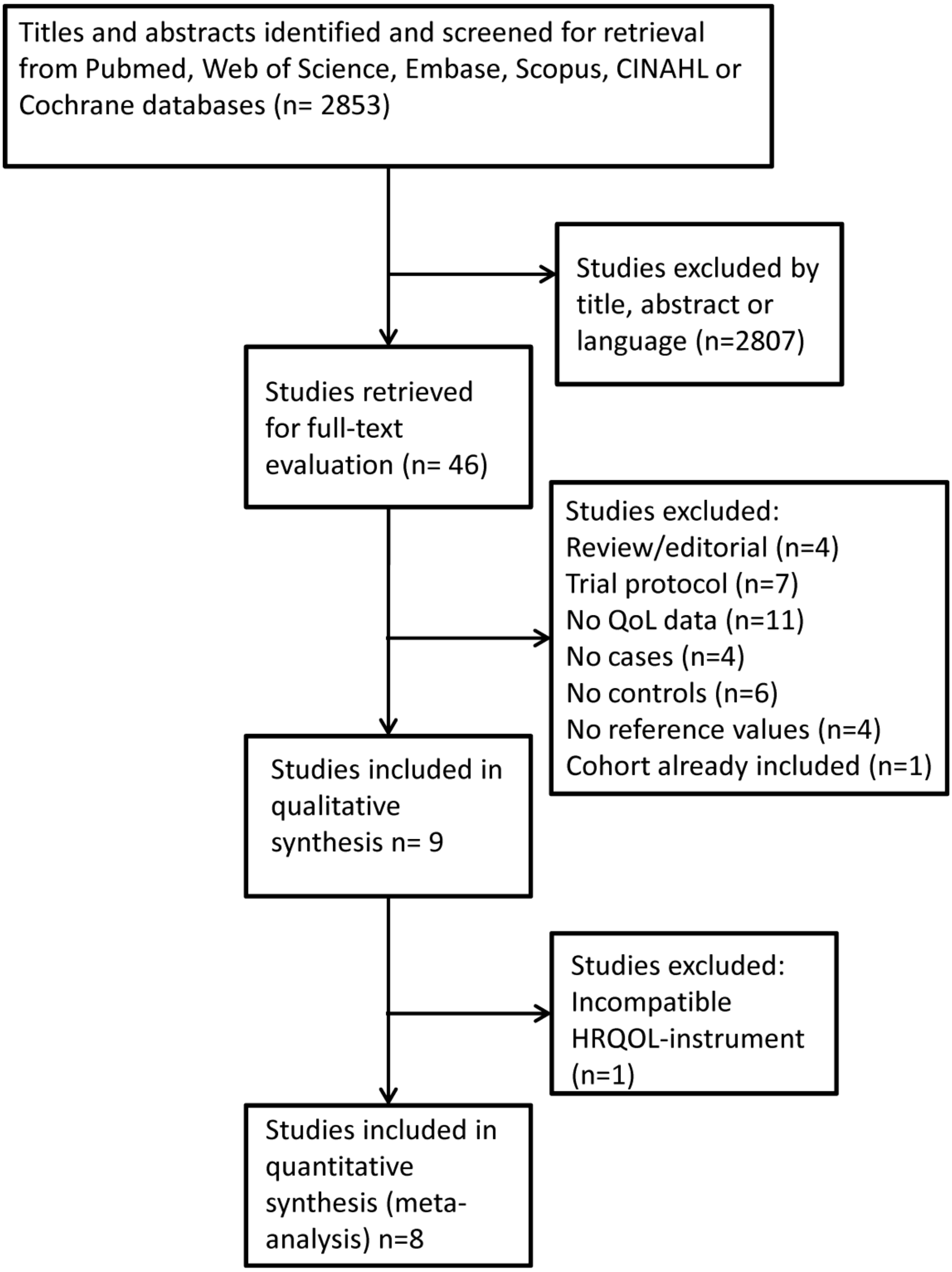
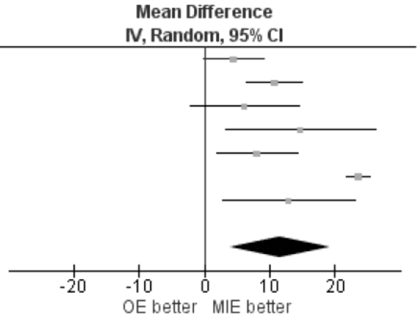

Titles and abstracts identified and screened for retrieval from Pubmed, Web of Science, Embase, Scopus, CINAHL or Cochrane databases (n= 2853)
Studies excluded by title, abstract or language (n=2807)
Studies retrieved for full-text evaluation (n= 46)
Studies excluded:
Review/editorial (n=4)
Trial protocol (n=7)
No QoL data (n=11)
No cases (n=4)
No controls (n=6)
No reference values (n=4)
Cohort already included (n=1)
Studies included in qualitative synthesis n= 9
Studies excluded:
Incompatible HRQOL-instrument (n=1)
Studies included in quantitative synthesis (meta-analysis) n=8

Figure 2

A) Global HRQOL

Study or Subgroup	MIE Mean	MIE SD	MIE Total	OE Mean	OE SD	OE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	60	23	373	55.5	21.4	110	15.3%	4.50 [-0.13, 9.13]
Hong 2013	82.1	10.5	51	71.3	11.9	55	15.5%	10.80 [6.53, 15.07]
Nafteux 2011	62.5	23	52	56.3	23	69	13.7%	6.20 [-2.08, 14.48]
Shen 2015	71.1	23	33	56.4	23	29	12.0%	14.70 [3.23, 26.17]
Wang 2010	71	11.9	27	62.9	12.1	29	14.7%	8.10 [1.81, 14.39]
Wang 2015	66.2	13	444	42.6	14.1	444	16.1%	23.60 [21.82, 25.38]
Zeng 2012	62.9	23	30	49.9	23.2	60	12.7%	13.00 [2.89, 23.11]
Total (95% CI)			**1010**			**796**	**100.0%**	**11.61 [3.84, 19.39]**

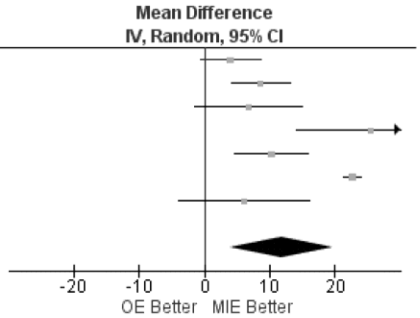
Heterogeneity: Tau² = 96.95; Chi² = 99.58, df = 6 (P < 0.00001); I² = 94%
Test for overall effect: Z = 2.93 (P = 0.003)



B) Physical function

Study or Subgroup	MIE Mean	MIE SD	MIE Total	OE Mean	OE SD	OE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	78	21.8	373	74	21.6	110	15.3%	4.00 [-0.60, 8.60]
Hong 2013	84.7	11.4	51	76.1	12.3	55	15.3%	8.60 [4.09, 13.11]
Nafteux 2011	76.7	23	52	69.9	23	69	13.7%	6.80 [-1.48, 15.08]
Shen 2015	89.6	23	33	64.1	23	29	12.1%	25.50 [14.03, 36.97]
Wang 2010	82.2	10.3	27	71.9	11.3	29	14.9%	10.30 [4.64, 15.96]
Wang 2015	72	11.1	444	49.3	9.7	444	16.0%	22.70 [21.33, 24.07]
Zeng 2012	94.3	23	30	88.2	23.1	60	12.8%	6.10 [-3.99, 16.19]
Total (95% CI)			**1010**			**796**	**100.0%**	**11.88 [3.92, 19.84]**

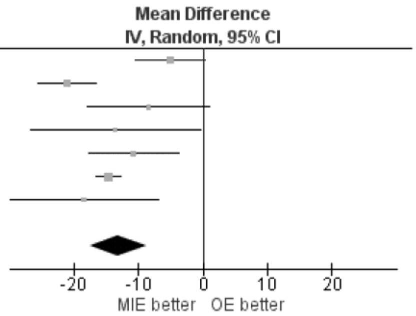
Heterogeneity: Tau² = 102.58; Chi² = 113.84, df = 6 (P < 0.00001); I² = 95%
Test for overall effect: Z = 2.92 (P = 0.003)



C) Fatigue

Study or Subgroup	MIE Mean	MIE SD	MIE Total	OE Mean	OE SD	OE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	41	25.3	373	46	25.5	110	17.1%	-5.00 [-10.41, 0.41]
Hong 2013	28.7	12.1	51	49.8	11.3	55	18.7%	-21.10 [-25.57, -16.63]
Nafteux 2011	34.8	26	52	43.2	27	69	11.1%	-8.40 [-17.91, 1.11]
Shen 2015	41.7	26	33	55.3	27	29	7.5%	-13.60 [-26.84, -0.36]
Wang 2010	40	13	27	50.7	13.9	29	14.5%	-10.70 [-17.75, -3.65]
Wang 2015	40.3	14.3	444	54.9	15	444	22.1%	-14.60 [-16.53, -12.67]
Zeng 2012	26.1	26	30	44.5	27.2	60	8.9%	-18.40 [-29.97, -6.83]
Total (95% CI)			**1010**			**796**	**100.0%**	**-13.18 [-17.59, -8.76]**

Heterogeneity: Tau² = 22.00; Chi² = 23.31, df = 6 (P = 0.0007); I² = 74%
Test for overall effect: Z = 5.85 (P < 0.00001)



D) Pain

Study or Subgroup	MIE Mean	MIE SD	MIE Total	OE Mean	OE SD	OE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	24	26.5	373	30	25.3	110	17.0%	-6.00 [-11.44, -0.56]
Hong 2013	17.1	8.2	51	37.1	9.7	55	19.8%	-20.00 [-23.41, -16.59]
Nafteux 2011	22.3	23	52	32.9	28	69	11.9%	-10.60 [-19.70, -1.50]
Shen 2015	35.7	26.5	33	63.9	28	29	7.5%	-28.20 [-41.82, -14.58]
Wang 2010	20.4	10.7	27	33.3	19.4	29	13.1%	-12.90 [-21.03, -4.77]
Wang 2015	30.8	10.7	444	50.7	12.5	444	21.7%	-19.90 [-21.43, -18.37]
Zeng 2012	9.9	26.5	30	26.3	27.8	60	9.0%	-16.40 [-28.21, -4.59]
Total (95% CI)			**1010**			**796**	**100.0%**	**-15.85 [-20.45, -11.24]**

Heterogeneity: Tau² = 24.86; Chi² = 30.86, df = 6 (P < 0.0001); I² = 81%
Test for overall effect: Z = 6.75 (P < 0.00001)

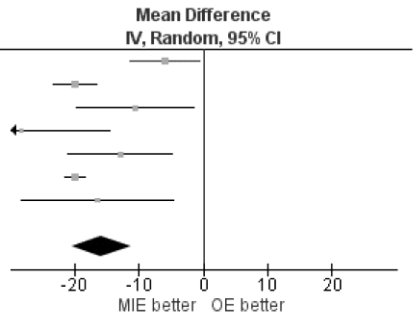
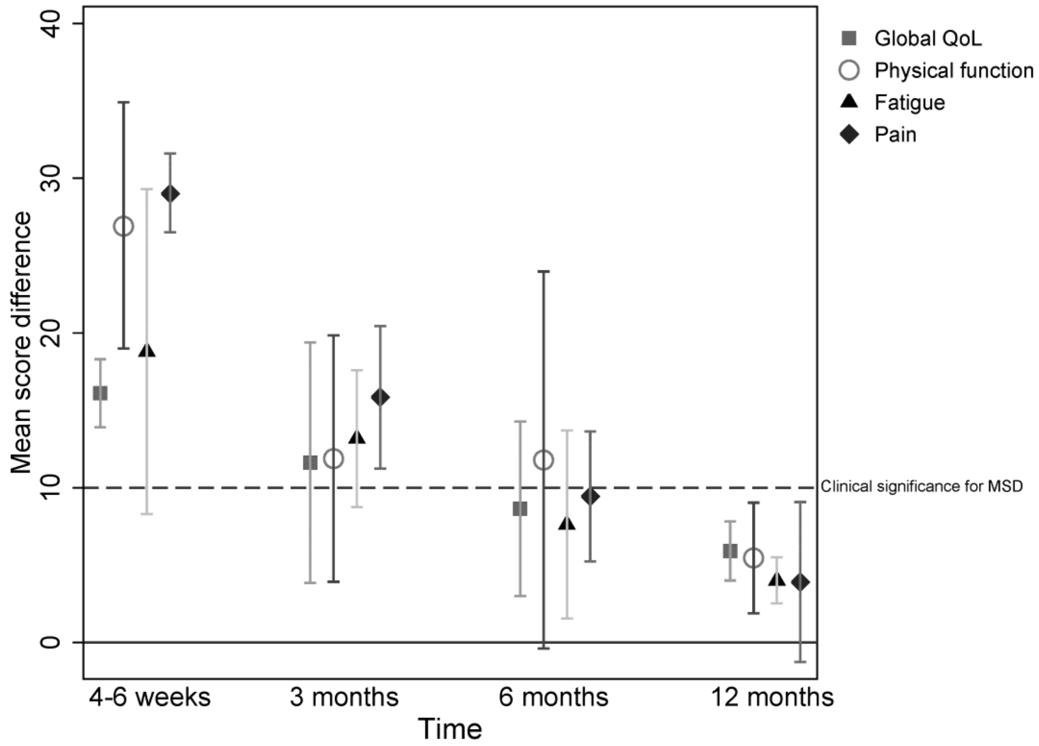


Figure 3



Global QoL
Physical function
Fatigue
Pain
Mean score difference
0
10
20
30
40
Clinical significance for MSD
4-6 weeks
3 months
6 months
12 months
Time

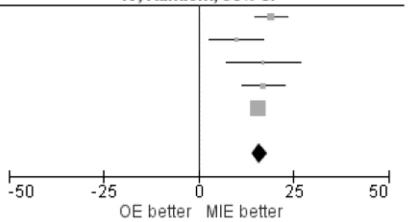
Suppl. Figure 1

A) Global HRQOL

	MIE			OE				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Hong 2013	62.6	11.8	51	43.4	10.1	55	19.7%	19.20 [15.00, 23.40]
Maas 2015	61	18	59	51	21	56	8.2%	10.00 [2.83, 17.17]
Shen 2015	69.1	18	33	52.1	21	29	4.6%	17.00 [7.20, 26.80]
Wang 2010	59.3	11.4	27	42.2	9.9	29	12.5%	17.10 [11.49, 22.71]
Wang 2015	55.7	12.4	444	40.1	9.2	444	54.9%	15.60 [14.16, 17.04]
Total (95% CI)			**614**			**613**	**100.0%**	**16.10 [13.91, 18.29]**

Heterogeneity: Tau² = 1.73; Chi² = 5.37, df = 4 (P = 0.25); I² = 25%
Test for overall effect: Z = 14.44 (P < 0.00001)

Mean Difference
IV, Random, 95% CI

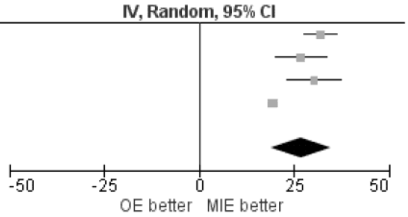


B) Physical function

	MIE			OE				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Hong 2013	72.9	10.4	51	40.8	12.1	55	26.0%	32.10 [27.81, 36.39]
Shen 2015	83	11.9	33	56.1	15.2	29	23.2%	26.90 [20.04, 33.76]
Wang 2010	75.7	11.9	27	45.4	15.2	29	22.9%	30.30 [23.18, 37.42]
Wang 2015	66.8	9.2	444	47.4	8.1	444	27.9%	19.40 [18.26, 20.54]
Total (95% CI)			**555**			**557**	**100.0%**	**26.93 [19.01, 34.85]**

Heterogeneity: Tau² = 58.13; Chi² = 42.26, df = 3 (P < 0.00001); I² = 93%
Test for overall effect: Z = 6.66 (P < 0.00001)

Mean Difference
IV, Random, 95% CI

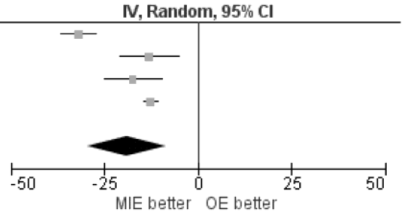


C) Fatigue

	MIE			OE				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Hong 2013	37.9	12.1	51	69.8	12.2	55	25.8%	-31.90 [-36.53, -27.27]
Shen 2015	55.3	14.2	33	68.3	16.7	29	23.7%	-13.00 [-20.77, -5.23]
Wang 2010	48.1	13.9	27	65.5	15.7	29	23.7%	-17.40 [-25.16, -9.64]
Wang 2015	44.2	14.2	444	56.8	16.7	444	26.9%	-12.60 [-14.64, -10.56]
Total (95% CI)			**555**			**557**	**100.0%**	**-18.81 [-29.34, -8.28]**

Heterogeneity: Tau² = 106.25; Chi² = 56.59, df = 3 (P < 0.00001); I² = 95%
Test for overall effect: Z = 3.50 (P = 0.0005)

Mean Difference
IV, Random, 95% CI

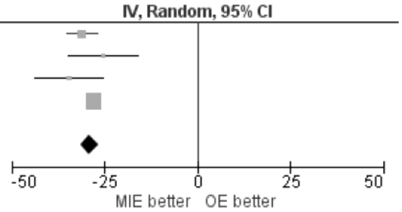


D) Pain

	MIE			OE				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Hong 2013	27.3	10.3	51	58.4	11.6	55	26.0%	-31.10 [-35.27, -26.93]
Shen 2015	57.7	13.8	33	83	22.2	29	6.8%	-25.30 [-34.65, -15.95]
Wang 2010	23.5	11.5	27	58	22.2	29	7.1%	-34.50 [-43.67, -25.33]
Wang 2015	34.8	13.8	444	62.7	13.9	444	60.1%	-27.90 [-29.72, -26.08]
Total (95% CI)			**555**			**557**	**100.0%**	**-29.02 [-31.56, -26.48]**

Heterogeneity: Tau² = 1.93; Chi² = 4.00, df = 3 (P = 0.26); I² = 25%
Test for overall effect: Z = 22.38 (P < 0.00001)

Mean Difference
IV, Random, 95% CI

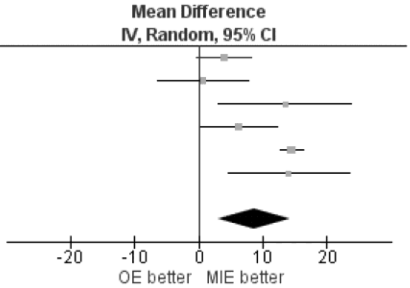


Suppl figure 2

A) Global HRQOL

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	64	22.7	364	60	19.1	106	19.2%	4.00 [-0.32, 8.32]
Nafteux 2011	64.6	22	48	63.9	13	60	16.2%	0.70 [-6.34, 7.74]
Shen 2015	74.8	22.7	33	61.3	19.1	29	12.5%	13.50 [3.09, 23.91]
Wang 2010	74.7	10.7	27	68.4	12.1	29	17.4%	6.30 [0.33, 12.27]
Wang 2015	68.5	11.6	444	54	15	444	21.2%	14.50 [12.74, 16.26]
Zeng 2012	85.6	22.7	30	71.5	19.3	60	13.4%	14.10 [4.62, 23.58]
Total (95% CI)			**946**			**728**	**100.0%**	**8.64 [2.99, 14.28]**

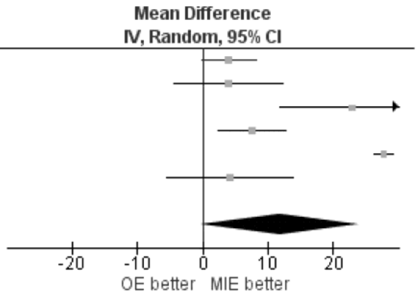
Heterogeneity: Tau² = 38.27; Chi² = 34.55, df = 5 (P < 0.00001); I² = 86%
Test for overall effect: Z = 3.00 (P = 0.003)



B) Physical function

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	81	20	364	77	19.3	106	17.4%	4.00 [-0.21, 8.21]
Nafteux 2011	81	22	48	77	22	60	16.4%	4.00 [-4.35, 12.35]
Shen 2015	94.1	22	33	71.2	22	29	15.5%	22.90 [11.92, 33.88]
Wang 2010	83.5	9	27	75.9	10.7	29	17.2%	7.60 [2.43, 12.77]
Wang 2015	77.1	12.1	444	49.3	9.7	444	17.7%	27.80 [26.36, 29.24]
Zeng 2012	96.9	22	30	92.7	22.2	60	15.9%	4.20 [-5.47, 13.87]
Total (95% CI)			**946**			**728**	**100.0%**	**11.79 [-0.40, 23.97]**

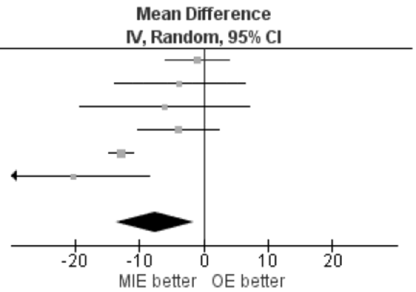
Heterogeneity: Tau² = 218.21; Chi² = 189.12, df = 5 (P < 0.00001); I² = 97%
Test for overall effect: Z = 1.90 (P = 0.06)



C) Fatigue

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	37	23.5	364	38	23.3	106	20.2%	-1.00 [-6.05, 4.05]
Nafteux 2011	31.7	27	48	35.4	26	60	14.3%	-3.70 [-13.78, 6.38]
Shen 2015	16.9	27	33	23	26	29	11.1%	-6.10 [-19.31, 7.11]
Wang 2010	23.7	12.4	27	27.6	11.5	29	18.8%	-3.90 [-10.18, 2.38]
Wang 2015	33.3	14	444	46.1	15.5	444	23.0%	-12.80 [-14.74, -10.86]
Zeng 2012	23.2	27	30	43.4	26	60	12.6%	-20.20 [-31.89, -8.51]
Total (95% CI)			**946**			**728**	**100.0%**	**-7.62 [-13.70, -1.55]**

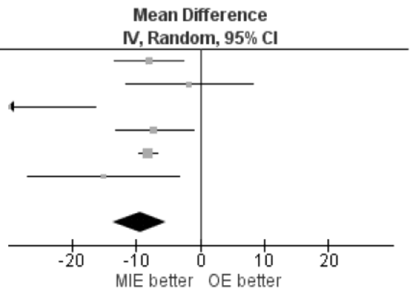
Heterogeneity: Tau² = 40.89; Chi² = 28.02, df = 5 (P < 0.0001); I² = 82%
Test for overall effect: Z = 2.46 (P = 0.01)



D) Pain

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	22	27.7	364	30	24.8	106	20.9%	-8.00 [-13.51, -2.49]
Nafteux 2011	20.5	26	48	22.2	26	60	11.6%	-1.70 [-11.57, 8.17]
Shen 2015	27	27.7	33	56.8	26	29	7.6%	-29.80 [-43.17, -16.43]
Wang 2010	19.2	11	27	26.4	12.2	29	19.4%	-7.20 [-13.28, -1.12]
Wang 2015	28.5	11.6	444	36.6	9.7	444	31.5%	-8.10 [-9.51, -6.69]
Zeng 2012	8.3	27.7	30	23.4	25.8	60	9.0%	-15.10 [-26.97, -3.23]
Total (95% CI)			**946**			**728**	**100.0%**	**-9.43 [-13.63, -5.24]**

Heterogeneity: Tau² = 14.00; Chi² = 13.11, df = 5 (P = 0.02); I² = 62%
Test for overall effect: Z = 4.41 (P < 0.0001)

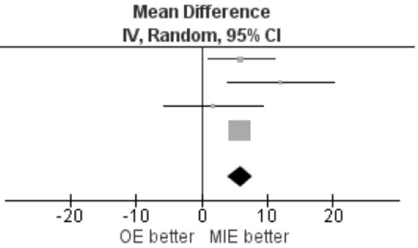


Suppl figure 3

A) Global HRQOL

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	69	22.4	316	63	20.2	80	13.0%	6.00 [0.93, 11.07]
Maas 2015	79	10	33	67	21	31	5.3%	12.00 [3.86, 20.14]
Nafteux 2011	67.7	20	48	65.9	21	67	6.1%	1.80 [-5.77, 9.37]
Wang 2015	70.5	13.3	444	64.7	10.4	444	75.7%	5.80 [4.23, 7.37]
Total (95% CI)			**841**			**622**	**100.0%**	**5.91 [4.00, 7.82]**

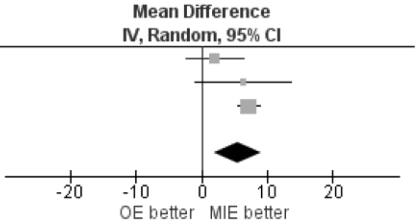
Heterogeneity: Tau² = 0.61; Chi² = 3.30, df = 3 (P = 0.35); I² = 9%
Test for overall effect: Z = 6.08 (P < 0.00001)



B) Physical function

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	85	20.8	316	83	17.4	80	30.8%	2.00 [-2.45, 6.45]
Nafteux 2011	82.4	18	48	76	22	67	16.9%	6.40 [-0.93, 13.73]
Wang 2015	78.8	12.3	444	71.6	13.3	444	52.2%	7.20 [5.51, 8.89]
Total (95% CI)			**808**			**591**	**100.0%**	**5.46 [1.89, 9.03]**

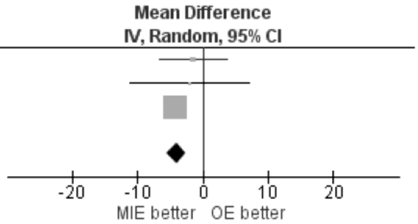
Heterogeneity: Tau² = 5.62; Chi² = 4.59, df = 2 (P = 0.10); I² = 56%
Test for overall effect: Z = 3.00 (P = 0.003)



C) Fatigue

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	30	24.4	316	31.5	20.5	80	8.1%	-1.50 [-6.74, 3.74]
Nafteux 2011	30.3	22	48	32.3	28	67	2.6%	-2.00 [-11.15, 7.15]
Wang 2015	26.9	13	444	31.2	10.8	444	89.3%	-4.30 [-5.87, -2.73]
Total (95% CI)			**808**			**591**	**100.0%**	**-4.01 [-5.50, -2.53]**

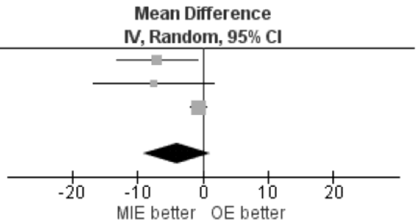
Heterogeneity: Tau² = 0.00; Chi² = 1.20, df = 2 (P = 0.55); I² = 0%
Test for overall effect: Z = 5.30 (P < 0.00001)



D) Pain

Study or Subgroup	OE Mean	OE SD	OE Total	MIE Mean	MIE SD	MIE Total	Weight	Mean Difference IV, Random, 95% CI
Barbour 2016	18	26.1	316	25	25.5	80	29.5%	-7.00 [-13.29, -0.71]
Nafteux 2011	17.4	24	48	24.9	26	67	19.7%	-7.50 [-16.71, 1.71]
Wang 2015	25.4	11.7	444	26.1	8	444	50.8%	-0.70 [-2.02, 0.62]
Total (95% CI)			**808**			**591**	**100.0%**	**-3.90 [-9.07, 1.27]**

Heterogeneity: Tau² = 13.23; Chi² = 5.59, df = 2 (P = 0.06); I² = 64%
Test for overall effect: Z = 1.48 (P = 0.14)



Suppl figure 4